

METHOD AND APPARATUS OF CONTROLLING NOISE LEVEL CALCULATIONS IN A CONFERENCING SYSTEM

Field of the Invention

5 This invention relates generally to audio conferencing systems, and more particularly to a method and apparatus for controlling noise level calculations in a conferencing system based on voice activity in a signal direction opposite to a that of a signal of interest.

Background of the Invention

10 In an audio conferencing system, whether full-duplex or half-duplex, it is useful to keep track of the noise level in both the incoming (line-in) and the outgoing direction (line-out). For reasons related to echo cancellation though, speech activity in the opposite direction of the signal of interest (that is, near-end
15 speech for line-in signal and far-end speech for line-out signal) may cause artificial fluctuations in the noise level that needs to be estimated. In other words, the absence of speech activity in the signal of interest does not guarantee that this portion of the signal represents the actual background noise of the signal of interest. Thus, where the signal of interest is the line-in signal, the echo canceller on the far-
20 end side either shuts down its transmit signal (in the case of a half-duplex device), or applies a "Non Linear Processor" (in the case of a full-duplex device) during speech activity in the received signal (near-end speech). This results in signal level variations in the 'line-in' signal during such near end speech activity which is misinterpreted as far end noise due to the absence of far-end speech. A similar
25 analysis applies to the noise level estimation of the line-out signal during far-end speech activity. In both cases, as indicated above, undesirable signal level variations result that may affect noise level estimations of the signal during speech (or tone) activity on the signal in the opposite direction.

Methods are well known in the art for tracking the level of the portions of a signal that are free of speech (or in-band tones) to perform noise level estimation. Thus, the prior art teaches the use of voice activity detection on a signal of interest to control noise level estimation on the signal. Example of such prior art systems are set forth in:

[1] "Noise signal prediction system". Joji Kane and Akira Nohara . US patent US5295225.

[2] "Noise suppression of acoustic signal in telephone set". Toshio Yoshida and Michitaka Sisido. US patent US5617472.

[3] "Method of detecting silence in a packetized voice stream". Franck Beaucoup. Mitel patent application #435.

None of the prior art, however, addresses the issue of noise level fluctuations due to speech activity on the signal in an opposite direction to the signal of interest. Consequently, the prior art systems discussed above may suffer from the aforementioned noise level fluctuations. The gravity of such consequences depends on the particular system; and in particular on how much tracking ability the application requires from the noise level estimation.

Summary of the Invention

According to the present invention, voice activity detection is applied to both the signal of interest and to the signal of opposite direction to the signal of interest itself in order to control the noise level calculation on the signal of interest. The method and apparatus of the present invention reduces the sensitivity of the noise level calculation to noise level fluctuations in the opposite direction signal, and therefore obtains a more accurate noise level estimation of the signal of interest.

Brief Description of the Drawings

A detailed description of the invention is set forth herein below, with reference to the drawings, in which:

Figures 1a and 1b are block diagrams of a line-in noise level estimator in accordance with first and second embodiments of the present invention;

Figures 2a and 2b are block diagrams of line-out noise level estimators in accordance with an alternative embodiment of the present invention; and

- 5 Figure 3 is a block diagram of line-in and line-out noise level estimator in accordance with the preferred embodiment.

Detailed Description of the Preferred Embodiment

- Turning to Figure 1a, a conferencing system is shown incorporating an
10 Acoustic Echo Canceller (AEC) block 1, as is well known in the prior art. In order to estimate and track the noise level of the incoming (line-in) signal, a Noise-Level-Estimator (NLE) block 2 is provided in the line-in signal path. As is also known in the prior art, the NLE block 2 is controlled by a Voice-Activity-Detector (VAD) block 3 on the line-in signal, so that only segments free of speech are used to update
15 the noise level calculation. However, in accordance with the present invention, another VAD block 5 on the line-out signal to ensure that the calculations in the NLE block 2 are also frozen during near-end speech. Preferably, the VAD block 3 includes a delay chosen to account for the network round-trip delay.

- 20 Instead of using first and second VAD blocks 3 and 5 after the AEC block 1, it is also possible to use only one VAD block 7 located on the line-out signal before the AEC block 1, as shown in Figure 1b. The VAD block 7 indicates both far-end (through the echo signal) and near-end speech and therefore freezes the calculations in the NLE block 2 in both cases.

25

In Figures 2a and 2b, equivalent block diagrams are provided to show the noise level estimation concepts of Figure 1a and 1b, respectively, applied to the case where the signal of interest is the line-out signal.

In some cases (e.g. energy/level based voice activity detection) the algorithm used in the VAD block itself requires an estimate of the noise level of the signal it operates on. In such cases, the symmetrical embodiment of Figure 3 can be used. Each NLE block 2A and 2B feeds its noise level estimates into the VAD blocks 9A and 9B, respectively, of the same signal, and is controlled by both VAD blocks (9A and 9B). More particularly, the VAD block outputs (i.e. 'voiced' / 'unvoiced' decisions) control the NLE blocks 2A and 2B. Whenever a controlling VAD's output indicates a 'voiced' segment in the signal the noise level calculation in a controlled NLE block is disabled (i.e. the NLE is 'frozen').

Variations and modifications of the invention are contemplated. Although the present invention applies specifically to audio signals, it can be used in applications where audio is not the only aspect of the system, for instance in combined audio-video conferencing systems. Also, the present invention applies not only to noise level calculations but more generally to the estimation of any characteristics of the background noise of a signal in any audio conferencing system.

All such alternative embodiments are believed to fall within the sphere and scope of the invention as defined by the appended claims.